

# Partitioned trace distances

Alexey E. Rastegin

*Department of Theoretical Physics, Irkutsk State University, Gagarin Bv. 20, Irkutsk 664003, Russia\**

New quantum distance is introduced as a half-sum of several singular values of difference between two density operators. This is, up to factor, the metric induced by so-called Ky Fan norm. The partitioned trace distances enjoy similar properties to the standard trace distance, including the unitary invariance, the strong convexity and the close relations to the classical distances. The partitioned distances cannot increase under quantum operations of certain kind including bistochastic maps. All the basic properties are re-formulated as majorization relations. Possible applications to quantum information processing are briefly discussed.

PACS numbers: 03.67.-a, 03.65.Ta, 02.10.Ud

Keywords: Trace distance, Ky Fan's norm, Metric properties, Strong convexity, Majorization, Monotonicity, Quantum operation

## I. INTRODUCTION

A quantification of closeness of quantum states is inevitable task for quantum information processing [1]. On the other hand, the spaces of quantum states are very interesting mathematical subjects [2]. If states are pure then comparison of them is not difficult. But all the real devices are exposed to noise. So the pure states used will eventually evolve to mixed states. This is not a unique reason for consideration of mixed states. As it is shown in [3], the cloning machine which can use any mixed states in symmetric space is very important in quantum computation. There are many ways to compare two mixed quantum states. It has been found that two measures, the trace distance [1, 2] and the fidelity [4–6], are widely useful in study of quantum information. For instance, the fidelity function is most frequently utilized as figure of merit for approximate cloning [7, 8], data compression schemes [9] and quantum broadcasting [10]. At the same time, the above measures are not able to describe the problem of state closeness in all respects. For example, the equality of fidelities for two pairs of density operators does not imply their unitary equivalence [11]. Recently, the sub-fidelity [12] and the super-fidelity [12, 13] have been studied. Some related measures were also used in the literature, such as the Bures distance [2], the Monge distance [14] and the sine distance [15]. In Ref. [16, 17] the Hilbert-Schmidt inner product was utilized as figure of merit.

On the whole, reasons for use of some distance measure are mainly provided by basic properties of the measure. These properties are usually related to the measurements, change under quantum operations and the convexity (concavity) in inputs. They must ensure convenient mathematical formalism for study of processes in quantum information. If some confidential information is encoded in quantum signals then any user will decode information by measurements in the final stage. That is, after quantum measurements he concludes from obtained data of measurement. The authors of [18] gave the scheme in which obtained statistics can then be used by the observer to reconstruct the tested state without measurement back-action. In view of this, a measure should be directly related to obtained data. Due to numerous scenarios, we rather need some collection of reliable measures complementing each other. As an example, the strength of cryptosystem B92 with respect to state-dependent cloning is revealed by relative error better than by global fidelity [19, 20]. So the notion of relative error allows to complete the picture of state-dependent cloning.

In the present work, we investigate a family of new distances between mixed quantum states. These distances are closely related to the trace distance which is obtained as particular case. Up to a factor, each distance is metric induced by the Ky Fan norm. Like the Schatten norms, the Ky Fan norms form a specially important class of unitarily invariant norms. In general, the unitarily invariant norms provide reasonable tools for obtaining distance bounds on quantum information processing [21, 22]. The distances induced by the Ky Fan norms enjoy the good features similar to properties of the standard trace distance. By construction, the described measures can naturally be called “partitioned trace distances.” In effect, together these distances minutely characterize a distinguishability of two quantum states via measurement statistics. It turns out that a set of all the partitioned distances gives more detailed distinctions of generated probability distributions than the standard trace distance. We also obtain other properties which lighten use of the partitioned trace distances.

---

\*Electronic address: rast@api.isu.ru

## II. DEFINITIONS

In this section, the definition of partitioned trace distances will be given. Let  $\mathcal{H}$  be  $d$ -dimensional Hilbert space. For any operator  $\mathbf{X}$  on  $\mathcal{H}$  the operator  $\mathbf{X}^\dagger \mathbf{X}$  is positive, that is  $\langle \psi | \mathbf{X}^\dagger \mathbf{X} | \psi \rangle \geq 0$  for all  $|\psi\rangle$ . The operator  $|\mathbf{X}|$  is defined as unique positive square root of  $\mathbf{X}^\dagger \mathbf{X}$ . The eigenvalues of  $|\mathbf{X}|$  counted with multiplicities are called the singular values  $s_j(\mathbf{X})$  of operator  $\mathbf{X}$  [23]. In the following, the singular values are arranged in decreasing order, that is  $s_1(\mathbf{X}) \geq s_2(\mathbf{X}) \geq \dots \geq s_d(\mathbf{X})$ . For  $k = 1, \dots, d$ , the Ky Fan  $k$ -norm is defined as [23]

$$\|\mathbf{X}\|_{(k)} := \sum_{j=1}^k s_j(\mathbf{X}) . \quad (2.1)$$

The norm  $\|\mathbf{X}\|_{(1)}$  is equal to the operator norm, and the norm  $\|\mathbf{X}\|_{(d)} = \text{tr}|\mathbf{X}|$  is the well-known trace norm [23]. Each operator norm induces some metric on quantum states. In particular, the trace distance between quantum states  $\rho$  and  $\varrho$  is defined by [1]

$$D(\rho, \varrho) := \frac{1}{2} \text{tr}|\rho - \varrho| \equiv \frac{1}{2} \|\rho - \varrho\|_{(d)} . \quad (2.2)$$

There is an alternative definition in terms of extremal properties of quantum operations [24]. The insertion of factor  $1/2$  is justified by analogy with the classical distance and by the bound  $D(\rho, \varrho) \leq 1$ . Let  $\{p_i\}$  and  $\{q_i\}$  be two probability distributions over the same index set. The  $L_1$ -distance (or Kolmogorov distance) is then defined as [1, 2]

$$\mathcal{D}(p_i, q_i) := \frac{1}{2} \sum_i |p_i - q_i| . \quad (2.3)$$

In view of the above reasons, it is natural to introduce new distances in the following way.

**Definition 2.1.** The  $k$ -th partitioned trace distance between density operator  $\rho$  and  $\varrho$  is defined by

$$D_k(\rho, \varrho) := \frac{1}{2} \|\rho - \varrho\|_{(k)} . \quad (2.4)$$

In simple case  $\dim(\mathcal{H}) = 2$ , the difference  $\rho - \varrho = (1/2)(\vec{u} - \vec{v}) \cdot \vec{\sigma}$  has eigenvalues  $\pm(1/2)|\vec{u} - \vec{v}|$  in terms of corresponding Bloch vectors [1]. So we obtain  $D_1(\rho, \varrho) = (1/4)|\vec{u} - \vec{v}|$  and  $D_2(\rho, \varrho) = (1/2)|\vec{u} - \vec{v}|$ . There are some clear properties of the introduced distances.

1. Bounds:  $0 \leq D_k(\rho, \varrho) \leq 1$ ;  $D_k(\rho, \varrho) = 0$  if and only if  $\rho = \varrho$ .
2. Symmetry:  $D_k(\rho, \varrho) = D_k(\varrho, \rho)$ .
3. Triangle inequality:  $D_k(\rho, \varrho) \leq D_k(\rho, \omega) + D_k(\omega, \varrho)$  for any three states  $\rho, \omega$  and  $\varrho$ .
4. If the states  $\rho$  and  $\varrho$  are pure then  $D_1(\rho, \varrho) = (1/2)D_d(\rho, \varrho)$  and  $D_k(\rho, \varrho) = D_d(\rho, \varrho)$  for  $k \geq 2$ .
5. Unitary invariance:  $D_k(\mathbf{U}\rho\mathbf{U}^\dagger, \mathbf{U}\varrho\mathbf{U}^\dagger) = D_k(\rho, \varrho)$  for any unitary operator  $\mathbf{U}$ .

In the following, we will essentially use the statement known as Ky Fan's maximum principle [23]. Let the eigenvalues  $\lambda_j$  of Hermitian operator  $\mathbf{X}$  be so arranged that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ . Then we have [25]

$$\sum_{j=1}^k \lambda_j = \max\{\text{tr}(\mathbf{P}\mathbf{X}) : \text{rank}(\mathbf{P}) = k\} , \quad (2.5)$$

where the maximization is over all projectors  $\mathbf{P}$  of rank  $k$ . Modifying the proof of theorem 1 of Ref. [25], this principle can be re-formulated as

$$\sum_{j=1}^k \lambda_j = \max\{\text{tr}(\Theta\mathbf{X}) : \mathbf{0} \leq \Theta \leq \mathbf{1}, \text{tr}(\Theta) = k\} , \quad (2.6)$$

where the maximum is taken over those positive operators  $\Theta$  with trace  $k$  that satisfy  $\Theta \leq \mathbf{1}$ . We do not enter into details here.

### III. CONVEXITY PROPERTIES

In this section, some convexity properties of partitioned trace distances will be established. It is known that extremal problems are of great importance in applied disciplines. The presence of convexity or concavity allows to simplify essentially a study of many extremal problems [26]. Recall that Hermitian operator  $(\rho - \varrho)$  can be represented in the form  $\rho - \varrho = R - T$ , where  $R$  and  $T$  are positive operators with orthogonal support spaces [1]. In linear algebra, this decomposition is usually referred to as Jordan's decomposition [23]. Let the  $\varkappa_r$ 's denote nonzero eigenvalues of  $R$ , and let the  $\tau_t$ 's denote nonzero eigenvalues of  $T$ . By the spectral decomposition of  $(\rho - \varrho)$ , we have

$$R = \sum_r \varkappa_r |r\rangle\langle r|, \quad (3.1)$$

$$T = \sum_t \tau_t |t\rangle\langle t|, \quad (3.2)$$

where the eigenvectors are normalized. If we mutually rearrange the values  $\varkappa_r$  and  $\tau_t$  in decreasing order then we obtain nonzero singular values  $s_j$  of  $(\rho - \varrho)$ . It is clear that  $|\rho - \varrho| = R + T$ . By Ky Fan's maximum principle,

$$2D_k(\rho, \varrho) = \max\{\text{tr}(\mathbf{P}|\rho - \varrho|) : \text{rank}(\mathbf{P}) \leq k\}. \quad (3.3)$$

Here the condition  $\text{rank}(\mathbf{P}) \leq k$  is correct due to  $D_l(\rho, \varrho) \leq D_k(\rho, \varrho)$  for  $l \leq k$ . Let us define the two specific subspaces for given  $k$ . The subspace  $\mathcal{L}_R$  is spanned by those  $|r\rangle$ 's that  $\varkappa_r \in \{s_1, \dots, s_k\}$ . The subspace  $\mathcal{L}_T$  is spanned by those  $|t\rangle$ 's that  $\tau_t \in \{s_1, \dots, s_k\}$ . The maximizing projector of minimal rank can be written as  $\mathbf{P} = \mathbf{P}_R + \mathbf{P}_T$ , where  $\mathbf{P}_R$  is projector onto  $\mathcal{L}_R$  and  $\mathbf{P}_T$  is projector onto  $\mathcal{L}_T$ . If zeros are contained in the set  $\{s_1, \dots, s_k\}$  then  $\text{rank}(\mathbf{P}) < k$ . For the above projector we have  $(\mathbf{P}_R - \mathbf{P}_T)(\rho - \varrho) = \mathbf{P}_R R + \mathbf{P}_T T \equiv \mathbf{P}|\rho - \varrho|$  and

$$2D_k(\rho, \varrho) = \text{tr}[(\mathbf{P}_R - \mathbf{P}_T)(\rho - \varrho)]. \quad (3.4)$$

**Theorem 3.1.** Let  $\{p_i\}$  and  $\{q_i\}$  be probability distributions over the same index set, and  $\rho_i$  and  $\varrho_i$  be density operators with the same index. Then

$$D_k\left(\sum_i p_i \rho_i, \sum_i q_i \varrho_i\right) \leq \sum_i p_i D_k(\rho_i, \varrho_i) + \mathcal{D}(p_i, q_i). \quad (3.5)$$

**Proof.** Let us put  $\rho = \sum_i p_i \rho_i$  and  $\varrho = \sum_i q_i \varrho_i$ . Using the Jordan decomposition of  $(\rho_i - \varrho_i)$  and the triangle inequality for real numbers, we see that

$$|\text{tr}[(\mathbf{P}_R - \mathbf{P}_T)(\rho_i - \varrho_i)]| \leq \text{tr}(\mathbf{P}|\rho_i - \varrho_i|). \quad (3.6)$$

Further,  $|\text{tr}[(\mathbf{P}_R - \mathbf{P}_T)\varrho_i]| \leq \text{tr}\varrho_i = 1$ . By these two inequalities and Eq. (3.4), the doubled left-hand side of Eq. (3.5) can be rewritten as

$$\begin{aligned} & \sum_i p_i \text{tr}[(\mathbf{P}_R - \mathbf{P}_T)(\rho_i - \varrho_i)] + \\ & \sum_i (p_i - q_i) \text{tr}[(\mathbf{P}_R - \mathbf{P}_T)\varrho_i] \\ & \leq \sum_i p_i \text{tr}(\mathbf{P}|\rho_i - \varrho_i|) + \sum_i |p_i - q_i| \\ & \leq \sum_i 2p_i D_k(\rho_i, \varrho_i) + 2\mathcal{D}(p_i, q_i), \end{aligned} \quad (3.7)$$

where the maximum principle has finally been used.  $\square$

For the whole trace distance the proved property is called "strong convexity" [1]. It must be stressed that the whole classical distance  $\mathcal{D}(p_i, q_i)$  is contained in Eq. (3.5). Indeed, the range of index  $i$  is independent of  $k$ . As a corollary of strong convexity, there is the joint convexity. Namely,

$$D_k\left(\sum_i p_i \rho_i, \sum_i p_i \varrho_i\right) \leq \sum_i p_i D_k(\rho_i, \varrho_i). \quad (3.8)$$

Substituting  $\varrho$  for all  $\varrho_i$ 's into Eq. (3.8), we obtain the convexity in the first input. That is,

$$D_k\left(\sum_i p_i \rho_i, \varrho\right) \leq \sum_i p_i D_k(\rho_i, \varrho). \quad (3.9)$$

Due to symmetry we also have convexity in the second input [1]. Using the reasons of this section, the triangle inequality can easily be obtained. We refrain from presenting the details here.

#### IV. RELATIONS WITH THE CLASSICALITY

Similar to the standard trace distance, the partitioned trace distances can closely be related with the corresponding classical distances. We shall now introduce the partitioned classical distances between two probability distributions. By  $n$  denote the cardinality of distributions  $\{p_i\}$  and  $\{q_i\}$ . For  $k = 1, \dots, n$ , the  $k$ -th partitioned classical distances between  $\{p_i\}$  and  $\{q_i\}$  is defined by

$$\mathcal{D}_k^\downarrow(p_i, q_i) := \frac{1}{2} \sum_{i=1}^k |p_i - q_i|^\downarrow, \quad (4.1)$$

where the arrows down indicate that the absolute values are put in the decreasing order. The distance  $\mathcal{D}_n(p_i, q_i)$  is the standard  $L_1$ -distance. [Because the distance  $\mathcal{D}_n(p_i, q_i)$  contains all the differences  $|p_i - q_i|$ , we justly omit the arrow down.] Let us suppose that two density operators  $\rho$  and  $\varrho$  are commuting. So they are diagonal in the same basis  $\{|i\rangle\}$ , that is

$$\rho = \sum_i \mu_i |i\rangle\langle i|, \quad (4.2)$$

$$\varrho = \sum_i \nu_i |i\rangle\langle i|. \quad (4.3)$$

It is clear that operator  $\rho - \varrho = \sum_i (\mu_i - \nu_i) |i\rangle\langle i|$  has singular values  $|\mu_i - \nu_i|^\downarrow$ . Due to the definition of the partitioned trace distances and Eq. (4.1), we have

$$D_k(\rho, \varrho) = \mathcal{D}_k^\downarrow(\mu_i, \nu_i). \quad (4.4)$$

Thus, if two density operators commute then the  $k$ -th partitioned trace distance between them is equal to the  $k$ -th classical distance between their eigenvalues. A connection can also be posed in terms of probabilities generated by a quantum measurement. A generalized measurement is described by so-called "positive operator-valued measure" (POVM). Recall that POVM  $\{\mathbf{M}_m\}$  is a set of positive operators  $\mathbf{M}_m$  satisfying [1, 2]

$$\sum_m \mathbf{M}_m = \mathbf{1}, \quad (4.5)$$

where  $\mathbf{1}$  is the identity operator. In general, this approach allows to extract more information from a quantum system than the projective measurements. For the two density operators, the traces  $\text{tr}(\mathbf{M}_m \rho) \equiv p_m$  and  $\text{tr}(\mathbf{M}_m \varrho) \equiv q_m$  are the probabilities of obtaining a measurement outcome labeled by  $m$ .

**Theorem 4.1.** For arbitrary two density operators  $\rho$  and  $\varrho$ , there is

$$D_k(\rho, \varrho) = \max\{\mathcal{D}_k^\downarrow(p_m, q_m) : \text{tr}(\mathbf{M}_m) \leq 1\}, \quad (4.6)$$

where the maximum is taken over those POVMs that  $\text{tr}(\mathbf{M}_m) \leq 1$  for all the POVM elements.

**Proof.** Using the expressions of probabilities  $p_m$ ,  $q_m$  and  $\rho - \varrho = \mathbf{R} - \mathbf{T}$ , we write

$$\begin{aligned} 2\mathcal{D}_k^\downarrow(p_m, q_m) &= \sum_{m=1}^k |p_m - q_m|^\downarrow \\ &= \sum_{m=1}^k |\text{tr}[\mathbf{M}_m(\mathbf{R} - \mathbf{T})]|^\downarrow. \end{aligned} \quad (4.7)$$

Due to  $\text{tr}(\mathbf{M}_m \mathbf{R}) \geq 0$  and  $\text{tr}(\mathbf{M}_m \mathbf{T}) \geq 0$ , we have

$$\begin{aligned} |\text{tr}[\mathbf{M}_m(\mathbf{R} - \mathbf{T})]| &\leq \text{tr}[\mathbf{M}_m(\mathbf{R} + \mathbf{T})] \\ &= \text{tr}(\mathbf{M}_m |\rho - \varrho|). \end{aligned} \quad (4.8)$$

It follows from Eqs. (4.7) and (4.8) that

$$2\mathcal{D}_k^\downarrow(p_m, q_m) \leq \text{tr}(\Theta |\rho - \varrho|), \quad (4.9)$$

where we put  $\Theta = \sum_{m=1}^k \mathbf{M}_m^\downarrow$ . [Operators  $\mathbf{M}_m^\downarrow$  are POVM elements rearranged with respect to the decreasing order of numbers  $|p_m - q_m|^\downarrow$ .] Using the completeness relation (4.5) and  $\text{tr}(\mathbf{M}_m) \leq 1$ , we get  $\Theta \leq \mathbf{1}$  and  $\text{tr}(\Theta) \leq k$ . Due to Eq. (2.6), the right-hand side of Eq. (4.9) is less than or equal to the sum of  $k$  largest eigenvalues of  $|\rho - \varrho|$ . In other words,  $2\mathcal{D}_k^\downarrow(p_m, q_m) \leq 2D_k(\rho, \varrho)$ . Let us show that the inequality is saturated for some POVM. We take the

projector-valued measure  $\{|r\rangle\langle r|\} \cup \{|t\rangle\langle t|\}$ , when each element  $\mathbf{M}_m$  is either  $|r\rangle\langle r|$  or  $|t\rangle\langle t|$ . For this measurement the value

$$|p_m - q_m| = |\text{tr}[\mathbf{M}_m(\mathbf{R} - \mathbf{T})]|$$

is equal to either  $\kappa_r$  or  $\tau_t$ . So the numbers  $|p_m - q_m|^\downarrow$  are just singular values of  $(\rho - \varrho)$ . We see that the equality in Eq. (4.6) is reached by the above PVM.  $\square$

It should be stressed that Theorem 4.1 deals with POVMs whose elements obey  $\text{tr}(\mathbf{M}_m) \leq 1$ . At the same time, for the whole trace distance the maximum can be taken over arbitrary POVMs [1]. However, our restriction is hardly essential from the operational point of view. By the Davies theorem [27], for many tasks the optimal POVM can be built of elements of rank one. For such POVMs the statement of Theorem 4.1 holds. Indeed, if  $\text{rank}(\mathbf{M}_m) = 1$  then a single nonzero eigenvalue of element  $\mathbf{M}_m$  does not exceed one due to (4.5), whence  $\text{tr}(\mathbf{M}_m) \leq 1$ . In the following, we will discuss the result (4.6) in the context of quantum information processing.

## V. FORMULATION IN TERMS OF MAJORIZATION

In this section, we pose the above results within majorization relations. Elements of majorization theory are fruitfully used in the researches of quantum systems [28]. For example, the disorder criterion for separability has been found in terms of majorization relations [29]. We shall now recall the basic notions of the theory of majorization [30]. Let  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  be elements of real space  $\mathbb{R}^n$ . Let  $x^\downarrow$  be the vector obtained by rearranging the coordinates of  $x$  in the decreasing orders, that is

$$x_1^\downarrow \geq x_2^\downarrow \geq \dots \geq x_n^\downarrow. \quad (5.1)$$

We say that  $x$  is *weakly submajorized* by  $y$ , in symbols  $x \prec_w y$ , when [30]

$$\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow, \quad k = 1, \dots, n. \quad (5.2)$$

If the inequality is saturated for  $k = n$  then we say that  $x$  is *majorized* by  $y$  [23, 30]. In our analysis, components of real vectors are positive. Let integer  $n$  denote the cardinality of set  $\{\mathbf{M}_m\}$ , so that  $n$  is the total number of measurement outcomes. In the statement (4.6), the absolute values  $|p_m - q_m| = |\text{tr}(\mathbf{M}_m \rho) - \text{tr}(\mathbf{M}_m \varrho)|$  are summands in  $\mathcal{D}_k^\downarrow(p_m, q_m)$ , the singular values  $s_j(\rho - \varrho)$  are summands in  $D_k(\rho, \varrho)$ . Considering these summands as components of two real vectors, we append zeros so that  $|p - q|$  and  $s(\rho - \varrho)$  have the same dimension equal to  $\max\{n, d\}$ . Then the result (4.6) can be re-formulated as follows.

**Corollary 5.1.** If the condition  $\text{tr}(\mathbf{M}_m) \leq 1$  is fulfilled for all the POVM elements then  $|p - q|$  is weakly submajorized by  $s(\rho - \varrho)$ ,

$$|p - q| \prec_w s(\rho - \varrho). \quad (5.3)$$

If the equality  $\mathcal{D}_n(p_m, q_m) = D_d(\rho, \varrho)$  is extra valid for POVM  $\{\mathbf{M}_m\}$  then

$$|p - q| \prec s(\rho - \varrho), \quad (5.4)$$

id est  $|p - q|$  is majorized by  $s(\rho - \varrho)$ .

The relations (5.3) and (5.4) have the advantage of physical interpretation of singular values  $s_j(\rho - \varrho)$  in terms of distinctions of probability distributions. The relation between the standard trace distance and the  $L_1$ -distance is the well-known result in quantum information theory [1, 2]. Namely,

$$D_d(\rho, \varrho) = \max \left\{ \mathcal{D}_n(p_m, q_m) : \mathbf{0} \leq \mathbf{M}_m \leq \mathbf{1}, \sum_m \mathbf{M}_m = \mathbf{1} \right\}, \quad (5.5)$$

where the maximum is taken over all the POVM measurements. So, if two density operators are close in the standard trace distance, then any measurement performed on those states will give probability distributions close to each other. But they closeness is still characterized by one quantity solely. In this regard, the concept of partitioned trace distances provides sensitive and flexible tools for comparing mixed quantum states. Instead of one measure  $D_d(\rho, \varrho)$ , we now have a collection of  $d$  measures  $D_k(\rho, \varrho)$  and  $d$  singular values  $s_j(\rho - \varrho)$ . A closeness of two states are now described by system of  $d$  equations of the form (4.6) or, equivalently, by the majorization relation (5.3). Despite evident importance of the result (5.5), it does not give as much detailed information as provided by the relation (5.3).

For each pair of states  $\rho$  and  $\varrho$  we have the specified measurement such that the maximum in (4.6) is reached for all  $k = 1, \dots, d$  (then  $n = d$ ). As it follows from the proof of Theorem 4.1, this measurement is described by the projector-valued measure  $\{|r\rangle\langle r|\} \cup \{|t\rangle\langle t|\}$ . [The vectors  $|r\rangle$  and  $|t\rangle$  are defined in (3.1) and (3.2).] By simple calculation, we obtain

$$|p_j - q_j|^\downarrow = s_j(\rho - \varrho), \quad k = 1, \dots, d. \quad (5.6)$$

Thus, if the two probability distributions  $\{p_j\}$  and  $\{q_j\}$  are known then singular values  $s_j(\rho - \varrho)$  may be estimate due to (5.6). Here we have a possibility of measurement of partitioned trace distances in physical experiments. In principle, this might be a subject of separate investigation.

## VI. MONOTONICITY PROPERTIES

In this section we shall prove that the partitioned trace distances cannot increase under quantum operation of certain kind. The formalism of quantum operations provides a unified treatment of possible state change in quantum theory [1]. Quantum operations can be realized via programmable quantum processors [31]. Let  $\mathcal{H}_A$  and  $\mathcal{H}_B$  be finite-dimensional Hilbert spaces. We will consider a map  $\mathcal{E}$

$$\rho_A \rightarrow \rho_B := \frac{\mathcal{E}(\rho_A)}{\text{tr}_B[\mathcal{E}(\rho_A)]}, \quad (6.1)$$

where an input  $\rho_A$  is normalized state on  $\mathcal{H}_A$  and an output  $\rho_B$  is normalized state on  $\mathcal{H}_B$ . If the map  $\mathcal{E}$  describes physical process then it must be linear and completely positive [1, 2]. One demands that  $0 \leq \text{tr}_B[\mathcal{E}(\rho_A)] \leq 1$  for each input  $\rho_A$ . Then the map  $\mathcal{E}$  is a quantum operation with the input space  $\mathcal{H}_A$  and the output space  $\mathcal{H}_B$  [1, 2]. Each completely positive map can be written in the operator-sum representation. Namely, we have

$$\mathcal{E}(\rho_A) = \sum_m \mathbf{E}_m \rho_A \mathbf{E}_m^\dagger, \quad (6.2)$$

where operators  $\mathbf{E}_m$  map the input space  $\mathcal{H}_A$  to the output space  $\mathcal{H}_B$  [1, 2]. The normalization implies that

$$\sum_m \mathbf{E}_m^\dagger \mathbf{E}_m \leq \mathbf{1}_A, \quad (6.3)$$

where  $\mathbf{1}_A$  is the identity on  $\mathcal{H}_A$ . When physical process is deterministic,  $\text{tr}_B[\mathcal{E}(\rho_A)] = 1$  and the upper bound in Eq. (6.3) is saturated. Then a map is trace-preserving completely positive (TPCP). Such kind of operations is quantum analogue of classical stochastic maps [2]. A map  $\mathcal{E}$  is unital when

$$\mathcal{E}(\mathbf{1}_A) = \mathbf{1}_B, \quad (6.4)$$

where  $\mathbf{1}_B$  is the identity on the output space  $\mathcal{H}_B$ . As it follows from Eqs. (6.2) and (6.4), for unital operation

$$\sum_m \mathbf{E}_m \mathbf{E}_m^\dagger = \mathbf{1}_B. \quad (6.5)$$

Of course, the conditions (6.3) and (6.5) are independent from each other. For example, the depolarizing and phase damping channels are unital, while amplitude damping is not [1]. If TCPM-map is unital then it is called "bistochastic" [2]. This is quantum analogue of bistochastic matrix, which is a stochastic matrix that leaves the uniform probability vector invariant [2]. Such matrices can also be used for realization density matrix via uniform ensemble [32]. Both the conditions (6.3) and (6.5) are fulfilled for bistochastic map.

**Theorem 6.1.** If TCPM-map satisfies the condition

$$\sum_m \mathbf{E}_m \mathbf{E}_m^\dagger \leq \mathbf{1}_B, \quad (6.6)$$

then for arbitrary two normalized inputs  $\rho_A$  and  $\varrho_A$

$$D_k(\mathcal{E}(\rho_A), \mathcal{E}(\varrho_A)) \leq D_k(\rho_A, \varrho_A). \quad (6.7)$$

**Proof.** In the following,  $\rho_B \equiv \mathcal{E}(\rho_A)$  and  $\varrho_B \equiv \mathcal{E}(\varrho_A)$ . As it has been shown in Section 3, there exist two mutually orthogonal projectors  $\Pi_R$  and  $\Pi_T$  such that

$$\begin{aligned} 2D_k(\rho_B, \varrho_B) &= \text{tr}_B[(\Pi_R - \Pi_T)(\rho_B - \varrho_B)] \\ &= \text{tr}_A[(\Theta_R - \Theta_T)(\rho_A - \varrho_A)] \\ &\leq \text{tr}_A[(\Theta_R + \Theta_T)|\rho_A - \varrho_A|]. \end{aligned} \quad (6.8)$$

Here we use the operator-sum representation, the properties of the trace and direct analogue of Eq. (3.6). We also put two positive operators

$$\begin{Bmatrix} \Theta_R \\ \Theta_T \end{Bmatrix} = \sum_m \mathbf{E}_m^\dagger \begin{Bmatrix} \Pi_R \\ \Pi_T \end{Bmatrix} \mathbf{E}_m . \quad (6.9)$$

Due to the precondition (6.6) and  $\text{rank}(\Pi_R + \Pi_T) \leq k$ , the trace of positive operator  $\Theta = \Theta_R + \Theta_T$  satisfies  $\text{tr}_A(\Theta) \leq \text{tr}_B(\Pi_R + \Pi_T) \leq k$ . For any  $|\psi_A\rangle \in \mathcal{H}_A$  we have

$$\begin{aligned} \langle \psi_A | \Theta | \psi_A \rangle &= \sum_m \langle \psi_A | \mathbf{E}_m^\dagger (\Pi_R + \Pi_T) \mathbf{E}_m | \psi_A \rangle \\ &\leq \sum_m \langle \psi_A | \mathbf{E}_m^\dagger \mathbf{E}_m | \psi_A \rangle \leq \langle \psi_A | \psi_A \rangle \end{aligned} \quad (6.10)$$

by  $(\Pi_R + \Pi_T) \leq \mathbf{1}_B$  and Eq. (6.3). This implies that  $\Theta \leq \mathbf{1}_A$ . Using Eq. (2.6), we then see that the right-hand side of Eq. (6.8) does not exceed  $2D_k(\rho_A, \varrho_A)$ .  $\square$

Thus, if TPCP-map satisfies the condition (6.6) then it is contractive with respect to all the partitioned trace distances. In particular, a bistochastic map is contractive. If we allow states to be not normalized then the condition of preservation of the trace can be excluded. Due to Eq. (6.3) the inequality (6.10) remains valid. So the studied distances cannot increase under each quantum operation that obeys Eq. (6.6). The statement of Theorem 6.1 can be reformulated as a majorization relation. Namely, if TPCP-map obeys the condition (6.6) then

$$s(\rho_B - \varrho_B) \prec_w s(\rho_A - \varrho_A) . \quad (6.11)$$

Together with the relation (5.3), the majorization relation (6.11) allows to foresee some information on measurement statistics at the output if statistics at the input of quantum channel is known *a priori*.

It must be stressed that the whole trace distance cannot increase under *arbitrary* trace-preserving quantum operation [1]. For the partitioned trace distances this is not the case. In effect, the Hilbert-Schmidt distance is also not contractive generally [33]. Nevertheless, the class of operations under which the partitioned distances are monotonous is wide enough. This class contains both the basic transformations, namely the unitary evolution and the measurement. In the case of measurement, we have  $\mathcal{H}_B = \mathcal{H}_A$  and

$$\rho_B = \sum_m \mathbf{M}_m^{1/2} \rho_A \mathbf{M}_m^{1/2} . \quad (6.12)$$

Here the completeness relation implies both the trace preservation and the unitality. Thus, no basic physical transformations ever increase the partitioned trace distances.

## VII. CONCLUSION

To sum up, we see that the partitioned trace distances provide a kind of physical interpretation for singular values of difference between two density operator. Let us consider shortly possible applications of the above results in quantum information processing. First, many extensively used channels satisfy the condition (6.6). In particular, this is fulfilled by bistochastic maps, including the depolarizing channel and phase damping channels. For these channels we can use all the consequences of majorization relations (5.3) and (6.11). Second, we may utilize the partitioned trace distances in study of quantum channel given as “black box.” Here we choose a set of probe states which is sufficiently dense in the space of density operators. Putting these probe states into black box, we may estimate the partitioned distances between outputs. If an increase of some distance has been detected then the tested quantum channel certainly violates the condition (6.6). Of course, this is rough draft of *modus operandi* only.

In the present paper we introduce a family of new distances between mixed quantum states one of which is the trace distance. In view of their construction, these distances have naturally be called “partitioned trace distances.” Each distance is, up to factor, a metric induced by some Ky Fan’s norm. Several important properties of introduced distances have been established. The partitioned trace distances enjoy the metric properties, the unitary invariance, the strong convexity and the close relations to the corresponding classical distances. In addition, they do not increase under quantum operations of certain kind including the bistochastic maps. The partitioned trace distances provide convenient tools for comparing mixed quantum states. So the described family of distances can quite be utilized as figure of merit for quantum information processing.

## Acknowledgments

This work was supported in a part by the Ministry of Education and Science of the Russian Federation under grants no. 2.2.1.1/1483, 2.1.1/1539. The present author is grateful to anonymous referees for helpful remarks.

- 
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
  - [2] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
  - [3] H. Fan, *Quantum cloning of mixed states in symmetric subspace*, Phys. Rev. A **68**, 052301 (2003).
  - [4] A. Uhlmann, *The transition probability in the state space of a \*-algebra*, Rep. Math. Phys. **9**, pp. 273–279 (1976).
  - [5] P. M. Alberti and A. Uhlmann, *Stochastic linear maps and transition probability*, Lett. Math. Phys. **7**, pp. 107–112 (1983).
  - [6] R. Jozsa, *Fidelity for mixed quantum states*, J. Mod. Opt. **41**, 2315–2323 (1994).
  - [7] A. E. Rastegin, *Upper bound on the global fidelity for mixed-state cloning*, Phys. Rev. A **68**, 012305 (2003).
  - [8] A. E. Rastegin, *Global-fidelity limits of state-dependent cloning of mixed states*, Phys. Rev. A **68**, 032303 (2003).
  - [9] I. Bjelaković and A. Szkola, *The data compression theorem for ergodic quantum information sources*, Quantum Information Processing **4**, 49–63 (2005).
  - [10] V. Giovannetti and A. S. Holevo, *Quantum shared broadcasting*, Quantum Information Processing **7**, 55–69 (2008).
  - [11] A. Uhlmann, *On "partial" fidelities*, Rep. Math. Phys. **45**, 407–418 (2000).
  - [12] J. A. Miszczak, Z. Puchała, P. Horodecki, A. Uhlmann and K. Życzkowski, *Sub- and super-fidelity as bounds for quantum fidelity*, Quantum Information & Computation **9**, pp. 0103–0130 (2009).
  - [13] P. E. M. F. Mendonça, R. d. J. Napolitano, M. A. Marchioli, C. J. Foster and Y.-C. Liang, *Alternative fidelity measure between quantum states*, Phys. Rev. A **78**, 052330 (2008).
  - [14] K. Życzkowski and W. Słomczyński, *The Monge metric on the sphere and geometry of quantum states*, J. Phys. A: Math. Gen. **34**, 6689–6722 (2001).
  - [15] A. E. Rastegin, *Sine distance for quantum states*, e-print arXiv:quant-ph/0602112v1 (2006).
  - [16] S. Adhikari, A. K. Pati, I. Chakrabarty and B. S. Choudhury, *Hybrid quantum cloning machine*, Quantum Information Processing **6**, 197–219 (2008).
  - [17] P. E. M. F. Mendonça, A. Gilchrist and A. C. Doherty, *Optimal tracking for pairs of qubit states*, Phys. Rev. A **78**, 012319 (2008).
  - [18] D. D'Alessandro and R. Romano, *Further results on the observability of quantum systems under general measurement*, Quantum Information Processing **5**, 139–160 (2006).
  - [19] A. E. Rastegin, *Relative error of state-dependent cloning*, Phys. Rev. A **66**, 042304 (2002).
  - [20] A. E. Rastegin, *A lower bound on the relative error of mixed-state cloning and related operations*, J. Opt. B: Quantum Semiclassical Opt. **5**, S647–S650 (2003).
  - [21] D. A. Lidar, P. Zanardi and K. Khodjasteh, *Distance bounds on quantum dynamics*, Phys. Rev. A **78**, 012308 (2008).
  - [22] K. Khodjasteh and D. A. Lidar, *Rigorous bounds on the performance of a hybrid dynamical–decoupling quantum–computing scheme*, Phys. Rev. A **78**, 012355 (2008).
  - [23] R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).
  - [24] A. E. Rastegin, *Trace distance from the viewpoint of quantum operation techniques*, J. Phys. A: Math. Theor. **40**, 9533–9549 (2007).
  - [25] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations. I*, Proc. Nat. Acad. Sci. U.S.A. **35**, 652–655 (1949).
  - [26] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1969).
  - [27] E. B. Davies, *Information and quantum measurement*, IEEE Trans. Inf. Theory **IT-24**, 596–199 (1978).
  - [28] M. A. Nielsen and G. Vidal, *Majorization and the interconversion of bipartite states*, Quantum Information & Computation **1**, 76–93 (2001).
  - [29] M. A. Nielsen and J. Kempe, *Separable states are more disordered globally than locally*, Phys. Rev. Lett. **86**, 5184–5187 (2001).
  - [30] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, 1979).
  - [31] V. Bužek, M. Hillery, M. Ziman and M. Roško, *Programmable quantum processors*, Quantum Information Processing **5**, 313–420 (2006).
  - [32] I. Bengtsson and A. Ericsson, *How to mix a density matrix*, Phys. Rev. A **67**, 012107 (2003).
  - [33] X. Wang and S. G. Schirmer, *On the contractivity of the Hilbert–Schmidt distance under open system dynamics*, e-print arXiv:0901.4547v2 [quant-ph] (2009).